



**VNIVERSIDAD  
D SALAMANCA**

**Máster Interuniversitario en**

**Lógica y Filosofía de la Ciencia**

TRABAJO FIN DE MÁSTER – CURSO 2013/2014

**FENOMENOLOGÍA HUSSERLIANA E  
INTELIGENCIA ARTIFICIAL EN  
HUBERT L. DREYFUS**

Ricardo Mejía Fernández

Director:

**Juan José Acero Fernández**

24 de junio de 2014

***Agradecimientos:***

*A Hubert L. Dreyfus, por prestarse a resolver mis dudas  
y a Juan José Acero, al proponerme este apasionante tema.*

## **TABLA DE MATERIAS**

### **Introducción**

- 1. Dreyfus: cartógrafo de la Inteligencia Artificial**
- 2. Un cambio de rumbo en la Inteligencia Artificial con inspiración fenomenológica**
  - 2.1. El desempeño del cuerpo**
  - 2.2. El desempeño de la situación**
  - 2.3. El desempeño de las intenciones y necesidades**
- 3. La traslación de la fenomenología husserliana a la heideggeriana a favor de la Inteligencia Artificial**
- 4. Conclusiones críticas**
- 5. Bibliografía**

## Introducción

Demos introducción a nuestro TFM. Hubert L. Dreyfus, el autor substancial de este trabajo, es un filósofo de nacionalidad estadounidense, profesor ya jubilado de la Universidad de California, en Berkeley, que nació en 1929 en Terre Haute; y todavía sigue impartiendo docencia a través de seminarios y conferencias. Es muy conocido, no solamente a causa de su presencia mediática, sino por su prolífica obra y como intérprete de los fenomenólogos más relevantes en conexión con cuestiones científicas y tecnológicas, en especial de Edmund Husserl, a quien consagró su obra doctoral, y Martin Heidegger, de quien ha publicado importantes obras colectivas y comentarios de sus textos más sobresalientes<sup>1</sup>. En Estados Unidos, Dreyfus es muy reconocido por haber puesto en la misma vereda a fenomenólogos, tecnólogos y científicos. Su hermano, Stuart Dreyfus, que se ha dedicado profesionalmente a la informática, ha auxiliado a nuestro autor en el descubrimiento de las aseveraciones y problemas de estas disciplinas, entre los que se encuentra aquellos de la Inteligencia Artificial.

Huelga decir que Dreyfus hizo durante décadas estancias de investigación tanto en el Instituto Tecnológico de Massachusetts (MIT), como en centros universitarios superiores de Europa. Sólo para ejemplificar el reconocimiento que ha recibido Dreyfus por sus esfuerzos intelectuales, bástenos indicar que hay dos gruesos volúmenes, publicados en el mismo MIT, en homenaje a su *iter* académico: en el participan filósofos de la talla de John Searle, Charles Taylor, Sean Kelly, Daniel Andler, *et altri*. Este TFM tiene vocación de ser fiel al pensamiento de Dreyfus, puesto que hay una unidad en todas sus empresas intelectuales, a la vez que original. No es

---

<sup>1</sup> También sobre Michel Foucault, al editar la obra conjunta de 1983, junto a Paul Rabinow, *Michel Foucault: Beyond structuralism and hermeneutics*, en la cual podemos leer una entrevista al *normalien* parisino. Por desgracia no incluiremos esta obra dentro del contenido de la tesis, aunque recomendamos con mucho encarecimiento su lectura.

nuestra intención presentarlo todo de Dreyfus sino ajustarnos a los puntos concretos, sabiendo que este es un TFM de espacio muy limitado, que pertenecen a nuestro modesto tema. En primer lugar, digamos cuál es la hipótesis principal que aquí queremos contrastar. Así pues, nuestra hipótesis es que el proyecto tecnológico de la Inteligencia Artificial y la fenomenología de cariz husserliano son muy próximos en la obra de Dreyfus, dando nosotros un mayor espacio a la influencia de Dagfinn Føllesdal (1932-), filósofo de origen noruego que determinará la imagen, para algunos sesgada, que nuestro autor tiene de Husserl. La segunda hipótesis es que los logros empíricos, descriptivos y técnicos de la Inteligencia Artificial han ayudado a clarificar cuestiones fenomenológicas. La tercera hipótesis –que, a nuestro juicio, es más fuerte que la anterior– es que, la fenomenología de la Inteligencia Artificial de Dreyfus ha podido ser, *mutatis mutandi*, un acicate para que este proyecto despejase incógnitas, dogmatismos y puliese la manera en que enfocaba las planificaciones y objetivos de la IA.

Pero debemos avisar al lector o lectora que contrastar estas hipótesis es francamente complicado. Dreyfus no ha sistematizado su postura en un único libro, sino siempre a propósito de otros temas y autores. Intentaremos, por consiguiente, mostrar la relación mutua, así como lo más propio de nuestro autor. Siguiendo el estilo de tipo Harvard, la metodología de este TFM es a la vez diacrónica, al tratar en su devenir los temas dreyfusianos, y sincrónica, al verlos en la unidad de todo lo pensado por Dreyfus y la relación fenomenología-ciencia / fenomenología-tecnología. La estructura de este trabajo sigue un orden que une la diacronía con la sincronía, de modo que en el primer punto (1) empezaremos por el Dreyfus que despunta como cartógrafo de la Inteligencia Artificial, conociéndola de primera mano en el MIT en dos estadios diacrónicos.

En el segundo punto (2) nos proponemos manifestar cómo Dreyfus ha demandado un cambio de rumbo, con respecto a la manera incipiente de la IA, inspirándose en la fenomenología, no solamente de Husserl –de la que se distancia– sino de otros autores menos concentrados en el *cogito*. En el tercer punto (3), veremos por qué se aparta Dreyfus de Husserl a tenor de Dagfinn Føllesdal, autor que ocupa un lugar considerable en este trabajo, viendo aquellos puntos en los que noruego ha transmitido a nuestro autor la concepción fregeana de Husserl, esto es, de una mente que *transacciona* – permítasenos la expresión– sentidos y referencias. En el cuarto punto (4), finalizaremos el núcleo de nuestro TFM abordando como Dreyfus, asintiendo al peso e influencia de Husserl en la IA, realiza una traslación a la fenomenología heideggeriana, hasta el punto de preferir ésta por aquélla en la nueva IA. Acabaremos, en consecuencia, con unas muy condensadas conclusiones críticas (5) a todo este trabajo e incluiremos la bibliografía consultada como colofón.

## **1. Dreyfus: cartógrafo de la Inteligencia Artificial**

Son muy interesantes las fases de la Inteligencia Artificial que Dreyfus investigó de manera tan adelantada en sus años. Se podría decir, inclusive, que ha sido de los primeros en cartografiar, en calidad de filósofo, los dos estadios diacrónicos principales de cada una de sus articulaciones. El profesor de Berkeley lo hizo durante diez años desde 1957 hasta 1967. Este enorme trabajo se encuentra en *What computers can't do. A critique*, publicado en 1972<sup>2</sup>. En cuanto a la primera fase sobre la *simulación cognitiva*, al estadounidense le llamó poderosamente la atención todo lo

---

<sup>2</sup> Hay otras ediciones posteriores, las cuales pudiéramos disponer y contrastar en otro tipo de investigaciones en el que veamos, fase por fase, cada una de las apreciaciones de Dreyfus al respecto. En este TFM preferimos la primera debido a que nos da la ventaja de ser más resumida y esquemática en pro de la brevedad de un TFM.

concerniente a la *traducción del lenguaje*, la *resolución de problemas* y el *reconocimiento de patrones*. Uno de los primeros logros de los ordenadores ha sido poder traducir el lenguaje, mediante diccionarios confeccionados con el significado simbólico más unívoco posible.

Pero Dreyfus cree que la resolución de problemas y el reconocimiento de patrones es de lo más relevante, sobre todo gracias a los esfuerzos de Newell, Shaw y Simon en el Instituto de Tecnología Carnegie. Lo que no puede ser superado por los ordenadores es la habilidad humana para reconocer patrones en condiciones de creciente dificultad: de hecho, todo cambia según los patrones, los cuales no pueden ser siempre unívocos sino que, cuando los seres humanos obedecen a patrones, estos "*maybe skewed, incomplete, deformed, and embedded in noise*" (Dreyfus 1972, 40), es decir, la mente humana puede dar unidad a lo que está deformado en un contexto fragmentario, lo cual no puede ser realizado al mismo nivel por un ordenador. De hecho los requisitos para tal reconocimiento en los humanos "*maybe 'so fine and so numerous' that, even if they could be formalized, a search through a branching list of such traits would soon become unmanageable as new patterns for discrimination were added*" (Dreyfus 1972, *Ibíd.*).

La manera en que los seres humanos discriminamos los patrones escapa a una formalización total, porque algunos son verbales pero la mayoría de ellos es no-verbal y, ergo, no sujeto a manipulación lógico-simbólica directa. Dicho lo cual, "*the traits may depend upon external and internal context and are thus not amenable to context-free specification*" (Dreyfus 1972, *Ibíd.*), esto es, la manera en la que nosotros comprendemos los patrones no es acontextual y está supedita a múltiples variables vitales internas y externas. Los ordenadores no puede distinguir lo esencial de lo inesencial así como todo lo periférico a lo informacional de sus datos, mientras que el ser humano sí puede hacerlo. Nosotros podemos dar lo que Husserl llamaba

en *Ideen I* una “unidad de sentido”, como leemos en la referencia que hace Gómez Romero del pensador austríaco:

El fenomenólogo contempla toda realidad de la actitud natural como una «unidad de sentido» y, por consiguiente, el objeto adecuado de su investigación es descubrir en qué consiste y cómo se produce esa unidad de sentido; esto es lo que se denomina *fenomenología de la constitución*: ‘todas las unidades reales en sentido estricto son «unidades de sentido». Las unidades de sentido presuponen...una conciencia que dé sentido, que por su parte sea absoluta y no exista por obra de un dar sentido. (Gómez Romero 1995, 564).

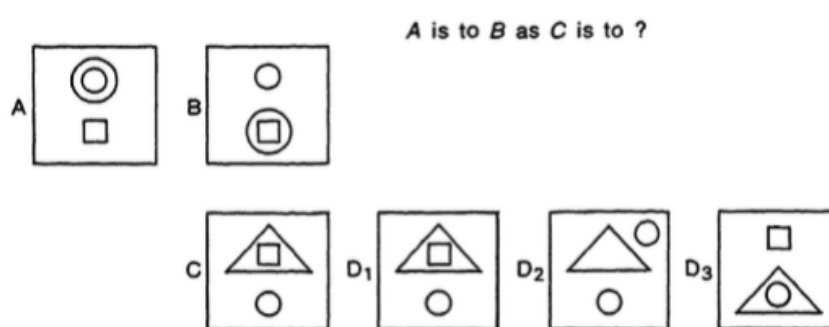
Es esta “actitud natural” del ser humano según Husserl que Dreyfus juzga que carecen las máquinas dichas inteligentes ya que, naturalmente, la unidad de sentido requiere previamente de una conciencia que es más que un software, es una conciencia que está inscrita y sumergida de raíz en el mundo de la vida, como se comprobó en *Krisis*. Por ello, la fenomenología de la constitución husserliana es tan diferente de la simulación cognitiva: cuando la conciencia da unidad de sentido a significados dispersos no simula nada ni se sobrepone a los contenidos del mundo de la vida. La conciencia humana no reconoce patrones sino que ella misma es naturalmente patrón noético y contenido hilético. ¿Qué es la actitud natural en Husserl? ¿Es un desdiseño de la naturaleza? Es más bien –y así queremos aclararlo en este TFM– un rechazo a una caracterización estrecha y cargada de ideología sobre la naturaleza. “Husserl la denomina ‘actitud naturalista’. Para la actitud naturalista, el mundo es mera naturaleza en el sentido de que se la contempla y estudia como una ‘*physis* carente de sentido’ (*bedeutungslose Physis*) donde únicamente imperan relaciones de causalidad. Sus objetos son considerados como ‘cosas en sí’ y, desde este punto de vista, los rasgos no naturales se muestran como ‘añadidos’ (*Annexe*) a esta cosa natural en sí” (Román Rabanaque 152). El proyecto tecnológico y científico de la



Inteligencia Artificial procede ese modo cuando privilegia la máquina sobre lo que circunda vitalmente a la máquina, es decir, cuando la máquina posee y manipula sus símbolos como cosas en sí y no como cosas vividas. Este es el falso naturalismo de la primera edición de la Inteligencia Artificial ya que, con su artificialidad, pretende deponer al ser humano en una superioridad que no es sino la total defección de no poder vivir las cosas en su sentido y significación unificadas en la raíz de la vida. Lo más contradictorio es que, pretendiendo ser artificiales y artefactuales, los ingenieros de la inteligencia artificial quieren emular la naturaleza para desfasarla, si bien lo que descartan en ella es quizá lo más complejo y valioso de cómo podemos habérnoslas con el mundo de formas más intuitivas, transversales y polisémicas.

Volviendo al terreno tecnológico, vemos que a partir de 1961 empezó un trabajo conducido por Marvin Minsky en el MIT sobre máquinas capaces de comprender el significado –sobre todo en su obra de 1964 *Semantic information processing*– justamente en los programas de idiomas extranjeros, si bien Dreyfus defiende que Minsky no fue tan crédulo en torno a lo que el proyecto tecnológico de la Inteligencia Artificial podía crear: "*a program now contains a few hundred facts. He estimates that 'a million facts would be for great intelligence'*" (Dreyfus 1972, 57). Por consiguiente y en cuanto al estadio diacrónico de *procesamiento semántico* de la información es donde más obstáculos tenemos, sobre todo en el test de la analogía de Evans, en donde se ve que los primeros ordenadores no podía estar expuestos a cuestiones más allá del procesamiento unívoco de información en maneras muy dispares a las directas orden-patrón-símbolo. Según las descripciones de Dreyfus (1972, *Ibíd.*) y dejándonos ilustrar por la siguiente figura, si nos basamos en las descripciones de D(A) y D(B), podríamos contar con muchas maneras en las que D(A) puede ser transformado en D(B). También muchas las maneras en las que las partes de A pueden ser puestas en correspondencia con las partes de C. Es decir

para pasar de A hasta D pueden hablar tantas posibilidades analógicas que no están disponibles a un tratamiento directo de A con D, la cuestión es que el ser humano maneja toda una compleja semántica que le permite descubrir las múltiples analogías de sus actos de habla y de su comprensión del mundo. En el gráfico siguiente se puede ver esta riqueza combinatoria en consonancia con las demás figuras geométricas que integran los cuadrados:



**Figura 1**

Encontramos más críticas de Dreyfus al proyecto de procesamiento semántico de la información: ya que, en el caso de resolución heurística de problemas, el programa informático está restringido a un dominio específico, un conjuntos de reglas y patrones, lo cual el mismo Minsky no pudo ignorar en los años del deslumbramiento del proyecto de la Inteligencia Artificial: "*Minsky criticizes the early programs for their lack of generality. 'Each program worked only on its restricted specialty, and there was no way to combine two different problem- solvers'* (Dreyfus 1972, 58). Lo curioso que sorprendió a Dreyfus y a nosotros es que, en vez de Minsky quejarse de la falta de contextualidad que permite la analogía, lo que hace es quejarse de la falta de generalidad: hemos visto en el gráfico sobre la analogía de Evans que precisamente lo difícil –imposible para muchos– es

que el ordenador pueda sumergirse en todo un juego de actos analógicos al igual que los seres humanos.

El Marvin Minsky posterior, principalmente en su libro *The Society of Mind*, llegó a modificar esta tendencia generalizante para pedir una mayor configuración y programación contextual de las máquinas inteligentes: "*It became clear that, in order to solve complicated problems, any machine of limited size must be able to reuse its agents in different ways in different contexts*" (Minsky 1986, 76). Incluso el científico une el contexto a la ambigüedad, ya que no se puede hacer una distinción tajante entre la expresión y el pensamiento. Siempre habrá condicionantes en el contexto de habla humana que nos darán ambigüedad: por eso no podemos dar una equivalencia total a los ordenadores y a las personas. Realmente y en lugar de ser un defecto, la ambigüedad del lenguaje que padece el ser humano es una competencia para adentrarse en las palabras que nombran el mundo. Así lo ha llegado a escribir Minsky: "*We can tolerate the ambiguity of words because we are already so competent at coping with the ambiguity of words*" (Minsky 1986, *Ibíd.*).

Por ello, lo que Dreyfus pedía mucho antes que Minsky –pudiera sorprender a más de uno que un filósofo exigía esto antes que un científico y provocó la mejora en este planteamiento– era, justamente, que todavía necesitamos saber "*the kind of heuristics we need to find heuristics, as well as what languages can readily describe them*" (Dreyfus 1972, 62). Mas, para nuestro autor, no debía ser tanto la aplicación general de reglas y su tratamiento simbólico el reto de aquella incipiente Inteligencia Artificial sino una *heuresis* de las variables informacionales y del lenguaje en toda su contextualidad. Dreyfus (1972, 63) deploraba que en aquellos años el proyecto de la Inteligencia Artificial sobresalía especialmente por su poco afán auto-crítico, su fe ciega en un progreso científico incierto e, inclusive, por la defectuosa

creatividad que se requería para hacer avanzar este campo que hoy incluimos en las ciencias cognitivas:

In fact, artificial intelligence is the least self-critical field on the scientific scene. There must be a reason why these intelligent men almost unanimously minimize or fail to recognize their difficulties, and continue dogmatically to assert their faith in progress. (Dreyfus 1972, 63)

Es así que Dreyfus ha tenido el mérito y el atrevimiento de haber propuesto alternativas a las aserciones tradicionales –las veremos en el próximo apartado– que dogmatizan a la Inteligencia Artificial y que se encubren en un optimismo naif, en especial las que (1) *biologizan* todo lo que la máquina inteligente opera, llegando a identificar –sobre todo en el ser humano– las operaciones mentales con las cerebrales –lo cual requiere un estudio reduccionista de las mismas mediante el predominante enfoque empírico de las neurociencias–, (2) las que *psicologizan*, las operaciones mentales de la máquina al hacerla semejante a los métodos introspeccionistas y de simulación cognitiva de la psicología experimental, (3) las que *epistemologizan* las operaciones inteligente de la máquina en comportamientos que se pueden formalizar en reglas y reproducir en cualquier otro soporte material o *hardware*, (4) las que ontologizan, al defender que el pensamiento se ha de descomponer en elementos primitivos o *bits*.

Pero, para Dreyfus, no basta lo que él bautiza como "*Platonic reduction*" (Dreyfus 1972, 143) del razonamiento en las máquinas inteligentes, abstrayendo todo lo cognoscible bajo la obediencia a reglas que formalizan los primitivos que constituyen su soporte físico y los *inputs* que vienen del exterior. Es necesario, por tanto, pensar otra Inteligencia Artificial que no esté presa de este platonismo trasnochado, es decir, que escoge el "*paradigm of logical intelligence*" y el "*essential attribute of rationality*" del

ser humano, como si otras dimensiones más afectivas, emocionales e intuitivas fuesen vedadas o imposibles en el tipo de máquinas inteligentes que se buscan. Por aquí se está moviendo hoy en día la Inteligencia Artificial y Dreyfus ha sido decisivo en esta *metanoia*. Véasmo, pues, en el próximo punto.

## **2. Un cambio de rumbo en la Inteligencia Artificial con inspiración fenomenológica**

A diferencia de los dos grandes estadios diacrónicos anteriores –el de la *simulación cognitiva* y el de la *semántica*– que, más bien, tienen como finalidad hacer que el desarrollo tecnológico de la Inteligencia Artificial se subordine a aserciones biólogos, psicólogos, epistemólogos y ontológicos, Hubert L. Dreyfus cree que es precisamente la antítesis a esto lo que debe de ser puesto en marcha. Ya que el proyecto clásico de la Inteligencia Artificial está viciado de raíz en sus presupuestos ideológicos – se inspira en un racionalismo reduccionista que se retrotrae a Platón–, lo que se debe hacer es replantearnos radicalmente que son falaces y filosofar acerca de un nuevo marco conceptual para la Inteligencia Artificial de los próximos años.

Aunque oigamos decir –en un complejo cientificista inmoderado– que los ordenadores son los que han de inspirar a las máquinas, Dreyfus cree que, aunque sea así, de cómo filosofemos sobre la realidad y sobre el ser humano dependerá dicha empresa tecnológica. De hecho, no es difícil dejar de ver en los estadios diacrónicos que en ellos en la pesquisa de dotar de inteligencia a las máquinas late una idea reduccionista del ser humano inteligente "*fascinated with the formalizable aspects of a subject*" (Dreyfus 1972, 145). Al otorgar el papel principal a la inteligencia formal en lugar de –digámoslo desde nuestra tradición iberoamericana– una "inteligencia

sentiente” (Zubiri 1980, 84). Pero Dreyfus, siendo extraño a nuestra tradición pero conocedor de la fenomenología continental francesa y alemana, clamaba por repensar la inteligencia en la IA desde la integralidad de lo humano, así distinguiéndola en lo que se pueda o no de lo humano:

Taking this suggestion to heart, we shall explore three areas necessarily neglected in CS and AI but which seem to underlie all intelligent behavior: the role of the body in organizing and unifying our experience of objects, the role of the situation in providing a background against which behavior can be orderly without being rulelike, and finally the role of human purposes and needs in organizing the situation so that objects are recognized as relevant and accessible. (Dreyfus 1972, 146)

A la conducta inteligente subyacen *tres desempeños*, que Dreyfus considera omitidos en la Inteligencia Artificial desde sus inicios: (1) el *desempeño del cuerpo* como estructurador de la inteligencia, pero no solamente como cuerpo biológico sino como cuerpo que es de *mi* experiencia objetual –lo cual es muy diferente de la noción con más dominancia en Husserl de una conciencia pura–, (2) el *desempeño de la situación* que se opone al psicologismo y al pensar que en la utopía de una formalización de todas las infinitas situaciones que se presenten –los términos y las reglas lógico-formales son finitas pero no así las circunstancias–, (3) el *desempeño de las intenciones y necesidades humanas* que dan sentido, pero también carácter pragmático, a la inteligencia humana.

Así estos tres desempeños integran una unidad de estructuración, trasfondo e intención, que ha de aguijar las nuevas rutas de la Inteligencia Artificial. Dreyfus nos descubre que los tres no han sido obviamente explotados por la fenomenología husserliana debido a su proximidad al racionalismo reduccionista que ha saturado lo que podría ser una reflexión más resuelta y pronto aceptar los aspectos menos desarrollados racional y formalmente

de la experiencia humana, así como su posibilitación natural. Tratemos cada uno de estos desempeños en los tres subapartados siguientes.

## **2.1. El desempeño del cuerpo**

Dreyfus cree que, por ejemplo el cuerpo se ha visto como algo más o menos relacionado con el pensamiento en la historia de la ciencia, si bien *"from Plato to Descartes has thought of the body as getting in the way of intelligence and reason, rather than being in any way indispensable for it"* (Dreyfus 1972, 147). Lo que sostiene el estadounidense es que el desempeño corporal es indispensable para la filosofía, para la tecnología y para la ciencia. En el caso de la Inteligencia Artificial el primer acento se puso en un mecanismo cibernético y no en lo que puede contribuir significativo un cuerpo vivo, en comparación a un sistema informático inerte. El mismo Descartes de los *Discursos* adelantó temas de robótica, pensando que las máquinas, algún día, podría realizar tareas de manera superior a los humanos. En la década de los 60, Herbert Simon pensaba (1965), en el mismo mecanicismo, que los ordenadores podrían realizar operaciones inteligentes que cualquier ser humano pudiese hacer.

Estamos en 2014 y eso no ha sucedido todavía, aunque contamos con Robots que tienen habilidades de mayor complejidad que el simple cálculo, la resolución de problemas previamente programados, etc. De la teoría de la percepción husserliana, Dreyfus (1972, 153) toma *"what Husserl calls the inner horizon"*. El horizonte interior, en el cual se unifican las diferentes percepciones y aspectos de un objeto, anticipa cualquier nueva percepción. Los expertos en el profesor de Friburgo explican que el horizonte interior *"comprises what we have recently been considering as 'absent' in a perception of an object: the further parts and aspects of the object itself that are not exhibited in a particular experience of the object, but are only*

'*empitly meant*'" (Smith 2003, 75). El problema es que, según el estadounidense, Husserl concibió este horizonte interior de manera internista y solipsista –con la conciencia trascendental–, sin atender debidamente al desempeño modificador del cuerpo, en especial del cerebro. En los comienzos de la Inteligencia Artificial no se podía comparar el procesador al cerebro ya que no recibe de manera pasiva la energía sino que la produce y, en sus cambios metabólicos se regenera de manera plástica.

No obstante, Dreyfus (1972, 160) era de la opinión de que "*Husserl is not free from the traditional intellectualist view*", de suerte que debe de ser complementado por fenomenologías de la corporalidad como la de Merleau-Ponty, el cual argumentaba que "*it is the body which confers the meanings discovered by Husserl*". Es el cuerpo el que permite el examen trascendental de Husserl, así como el análisis formal de los ingenieros de la Inteligencia Artificial. En lugar de la manera de proceder de los "*transcendental phenomenologists such as Husserl*" (Dreyfus 1972, 162), hemos de hacer que sobresalga "*our active, organically interconnected body*" en su propio funcionamiento. Lo mismo ha de ser advertido no solamente en el *software* y en los dispositivos fijos y no vivientes –el *hardware*– de la Inteligencia Artificial: el cuerpo es el garante de tres funciones inexistentes en los programas digitales: (1) el "horizonte interior" (*inner horizon*), del que hemos hablado y que anticipa lo que podemos percibir, (2) el "carácter global" (*global character*) que determina el significado, (3) la "transferabilidad" (*transferability*) entre todos los órganos del sistema corporal (Dreyfus 1972, 167). No puede haber conciencia pura ni horizonte trascendental, tampoco la mejor resolución de problemas o reconocimiento de patrones, sin este cuerpo que es horizonte, unificador y transmisor de las conductas inteligentes humanas.



## 2.2. El desempeño de la situación

En comparación con las aserciones epistemologistas y psicologistas, para Dreyfus la situación que sirve para no sólo para la programación en Inteligencia Artificial sino para que interpelan la supuesta conducta inteligente de las máquinas. El lenguaje informático o natural sin la situación no *funciona* y está desprovisto de lo que los lógicos han llamado verdad material del discurso. En la Inteligencia Artificial, mientras más lejanos estemos de la situación más inservibles e ineptos serán los programas, máquinas y robots que ingeniemos:

We must now turn directly to a description of the situation or context in order to give a fuller account of the unique way human beings are "in-the-world," and the special function this world serves in making orderly but nonrulelike behavior possible. (Dreyfus 1972, 168)

Este tan heideggeriano *ser-en-el-mundo* (*sein-in-der-Welt*) –en *Ser y tiempo* (Heidegger 2005)– es lo que nos faculta para una explicación más completa de la conducta epistémica en hombre y máquinas, las cuales se prestan a ambigüedad y se adaptan según la envolvencia circunstancial de un yo que no es un computador complejo de símbolos, un yo corporalizado en el mundo, es decir, expuesto a contextos de acción y de habla que no se pueden pre-decidir con total control. Por consiguiente, Dreyfus está instando a la tecnología y a la ciencia a aceptar el desafío de crear una inteligencia artificial cada vez más preparada a situaciones no programadas *ab initio*, es decir, posibles aplicaciones mucho más contextuales. Tenemos que evitar, con rotundidad, lo que el estadounidense denomina un "*computer-Cartesianism*" (Dreyfus 1972, 179) –detectado por él en el joven Minsky–, donde no sólo esté fuera de lugar el cuerpo, sino la situación del *cogito*.

En la actualidad cada vez se defiende menos el solipsismo metodológico – de una metodología atrincherada en los eventos y estados mentales– que, a partir de Descartes, se achaca por extensión a Husserl. Desde la misma filosofía del lenguaje, el texto, para ser significativo, ha de estar acompañado por su referencia y su con-texto: "*Wittgenstein constantly suggests that the analysis of a situation into facts and rules (which is where the traditional philosopher and the computer expert think they must begin) is itself only meaningful in some context and for some purpose*" (Dreyfus 1972, 175). El análisis compositivo de informáticos y de positivistas lógicos se subordina, para alcanzar su significatividad, al contexto y a las intenciones no solamente individuales sino sociales de la comunicación e interacción, sea ésta artificial o natural. En el caso de Husserl, en su etapa madura de *Krisis*, pudo defender el "mundo de la vida" (*Lebenswelt*) como subsuelo no preguntado (a)científicamente desde el paradigma moderno de la conciencia, el cual aqueja según Dreyfus a la Inteligencia Artificial:

[El mentado saber] no se pregunta nunca científicamente por el modo como el mundo de la vida funciona permanentemente como subsuelo, cómo sus múltiples valideces pre-lógicas son fundantes respecto de las lógicas, de las verdades teoréticas. Y tal vez es la científicidad que exige este mundo de la vida como tal y su universalidad, una científicidad apropiada, una que precisamente no es lógica-objetiva pero que como fundante último, respecto del valor, no resulta menor sino superior. (Husserl [1936] 2008, 166)

### **2.3. El desempeño de las intenciones y necesidades**

En conformidad con el pensamiento de Dreyfus, si por algo estamos en el mundo de la vida no es para ser simples tentáculos predeterminados suyos sino que el mundo de la vida cobra todo su significado y vivencia en nuestro

centro de experiencia: "*We are at home in the world and can find our way about in it because it is our world produced by us as the context of our pragmatic activity*" (Dreyfus 1972, 184). No es, por tanto, *el* mundo de la vida sino *nuestro* mundo tanto en las intenciones como en las necesidades individuales, colectivas y sociales.

Existen, por ello, diferencias entre los objetivos de una máquina y de una persona ya que nosotros no solamente nos adentramos en el mundo sino que nos apropiamos y lo configuramos según nuestro yo. Las máquinas son configuradas pero se les debe de añadir el reto de que puedan configurarse y configurar el mundo según las imprevisibilidades del mismo: esto es un nivel de exigencia que pone a la Inteligencia Artificial en la necesidad de ser mejorada de manera imparable. En el ser humano, los objetivos se miden según un sistema de valores no concisamente formalizado sino implícito, mientras que en las máquinas artificiales los comportamientos se evalúan según listas en relación de homeneidad simbólica y tablas formalizadas. Por eso, vemos cómo pueden prevalecer las listas enlazadas sobre los vectores, pero resulta que siempre hay un *quién* tras unas y otros, incluso en listas de enlace doble o circulares. La fenomenología se ha ocupado principalmente del método del examen de ese *Zentrum* de intenciones, si bien Dreyfus corrige al husserlianismo que la intencionalidad es corporalizada, situada y adaptada a las necesidades. Nuestra inteligencia es mucho más que algoritmos y almacenamiento/ordenamiento de datos (Cfr. Shaffer 1998).

En efecto, el ser humano está expuesto y es hábil para revoluciones culturales que germinan lo que no se contiene de manera seminal desde el inicio, mientras que los artefactos inteligentes precisan de un conjunto de reglas y de tablas bien pre-definidas para poder actuar y responder a problemas. La inteligencia humana es más que, como en la Inteligencia Artificial de la época de Dreyfus (1972, 190) "*the storage and processing of*

*data*”, es intuitiva, empática, desconfiada, tramposa, dada al juego etc. El ser humano no es solamente *homo faber* sino *homo ludens*. Debemos sostener, en consecuencia, la maleabilidad de la conducta humana siempre cambiante en función de lo intencionado y deseado: “*Man's nature is indeed so malleable that it may be on the point of changing again*” (Dreyfus 1972, 192). El profesor de Berkeley cita a Thomas Kuhn (1962, 17) exactamente para, mediante el concepto de paradigma, hacer ver que el ser humano no está conminado a procesamiento de reglas sino a una matriz disciplinar, que consta de un sistema de valores y de precomprensiones mucho más influyente en la heurística más aplicada: “*Paradigms may be prior to, more binding, and more complete than any set of rules for research that could be unequivocally abstracted from them*” (Dreyfus 1972, 1991).

Hemos de dejar de caracterizar la Inteligencia Artificial a partir de “*inflexible calculations of disembodied machines*” (Dreyfus 1972, *Ibíd.*) para proyectar si es o no una utopía la creación de máquinas flexibles capaces de adentrarse en el mundo de la vida y transformarlo en función de su intenciones y necesidades. La corporalidad, la situacionalidad y la intencionalidad-necesidad son de primera necesidad para la Inteligencia Artificial del porvenir, poniendo en evidencia, a su vez, las fallas de la pasada y actual.

### 3. Dagfinn Føllesdal: el maestro de Dreyfus en la fenomenología de Husserl

Toda interpretación que Dreyfus haga de Husserl está influida, para bien o para mal, por Dagfinn Føllesdal, su director doctoral en su formación harvardiana. Así pues, queremos indicar aquí que Føllesdal fue quien guió con más vigor a Hubert L. Dreyfus en la fenomenología de Husserl, siendo su director de doctorado en Boston. En *Husserl, Intentionality and Cognitive Science* (1982), encontramos al menos tres obras que Dreyfus recopila y donde se puede apreciar el gran peso que tiene Føllesdal en la propia imagen sobre la fenomenología y sobre Husserl que tiene Dreyfus. Aquí trataremos una en relación con nuestro tema. Laureado por Dreyfus como "*analytic philosopher and logician*" (Dreyfus y Hall 1982, 1), Føllesdal ha mantenido una postura fregeana a la hora de estudiar la intencionalidad en Husserl, en especial en una "*general theory of the contents of intentional states which accounted for the directedness of all mental activity*" (Dreyfus y Hall 1982, 2), ya que lo contaba era la generalidad de dichos contenidos intencionales sin perjuicio de la *direccionalidad* (*directedness*) de cómo la mente llega a ellos:

As Føllesdal explains more fully in the papers included in this volume, the phenomenological reduction is Husserl's way of describing the turning of attention away from both objects in the world and psychological activity to the mental contents which make possible the reference of each type of mental state to each type of object. (Dreyfus y Hall 1982, *Ibíd.*)

Esta es una visión de la reducción fenomenológica como acotadora de un significado fuera del mundo, de las cosas mismas y de los actos en los que está involucrada la psicología ordinaria. De hecho, el Dreyfus de los 80 se declaraba parte de una "*new generation of Husserl interpreters [that] has been working out what the noema is and what view of mind and reference*

*it implies*” (Dreyfus y Hall 1982, *Ibíd.*). La referencia en Frege y en Husserl se verá muy relacionada, aunque no menos confundida según los críticos. Michael K. Shim (2005), por ejemplo, ha criticado que Føllesdal, como luego será cortejado por Dreyfus, entienda el sentido y la referencia de modo simplemente conceptual, ya que se ha de contrastar la distinción husserliana *morphé* e *hylé*:

By reviewing a 35-year-old debate initiated by Dagfinn Føllesdal, I believe Husserlian phenomenology can be updated to offer an important contribution to this discussion. On my interpretation, what Husserl calls “hyletic data” may be read as that subjective quality of experience inarticulable as a propositional attitude – and, thus, hyletic data are non-conceptual. (Shim 2005, 209)

Lo hilético puede hacer alusión a lo que no es directamente concepción en una cualidad subjetiva previa a cualquier tipo de atribución. Pero esto es una crítica posterior. De hecho, el primer capítulo en la obra editada por Dreyfus está encargado a Føllesdal, obviamente –y no debe sorprendernos– sobre la percepción y constitución intención de objetos, todo ello de manera muy diferente a lo que un crítico como Shim (2005) y otros repensarán desde la materia hilética del conocimiento. Føllesdal (1982) investigó a Brentano y a Husserl a propósito de la percepción de estos objetos intencionales, lo cual será determinante también en la forma en la que ha trabajado Dreyfus desde aquellas décadas terciado por su maestro. La cuestión se hace más complicada si quisiésemos saber si estos objetos gozan de alguna suerte *direccionalidad*, como decíamos más arriba, y en el léxico de Føllesdal, una existencia o *inexistencia*:

Brentano, as you know, held that intentionality is characterized by a certain kind of directedness. We encounter problems when we try to characterize it, and you will remember that in his early writings Brentano simply said

that the directedness is characterized by there being some object which is always there, which the act is directed toward. Brentano's phrase is that the object "intentionally inexists" in our act. (Føllesdal 1982, 31)

En primer lugar, lo que puede calificarse como brentaniano es el tipo de direccionalidad de la mente como el *ser ahí* (*being there*) de cierto objeto pero, rozando la aparente contradicción, este ser ahí *inexiste* (*inextists*) en la intencionalidad mental. Esto es, en definitiva, lo que Føllesdal (1982, 32) llama un "dilema" (*dilemma*) ya que el ser ahí del objeto no es como el ser ahí de los objetos en el afuera mundanal sino que es un ser ahí intencional y solamente en esta forma podemos afirmar que inexisten toda vez que son intencionales. Una cosa es la objeto desde el afuera del mundo físico y el objeto intencionado en el adentro mental:

One argument consists in pointing out various differences between the real physical object and the object of thought. These are two different things, and in fact Brentano is not the first one who has emphasized their difference. We find the same observation made, for example, by Frege, and before him by Bolzano. One might, for example, hold that when a person is thinking about Pegasus, then really what he is thinking about is just his idea of Pegasus. Frege rejected this, saying that if that were the case then clearly with the same ground we could say that when somebody is thinking about the moon, what he is really thinking about is just his idea of the moon. And clearly an idea of the moon is quite a different thing from the moon itself. (Føllesdal 1982, *Ibíd.*).

Como acabamos de leer, una cosa es el "objeto físico real" y el "objeto de pensamiento". Si hablamos sobre lo pensado sobre Pegaso no hablamos sobre Pegaso, como objeto físico real, sino sobre el objeto de pensamiento. Lo mismo sucede con el pensamiento lunar, el cual por muy creativa que sea la idea no tiene que ver necesariamente con la objetividad física de la luna. Se tendría que contrastar el concepto con la cosa, o en el linaje

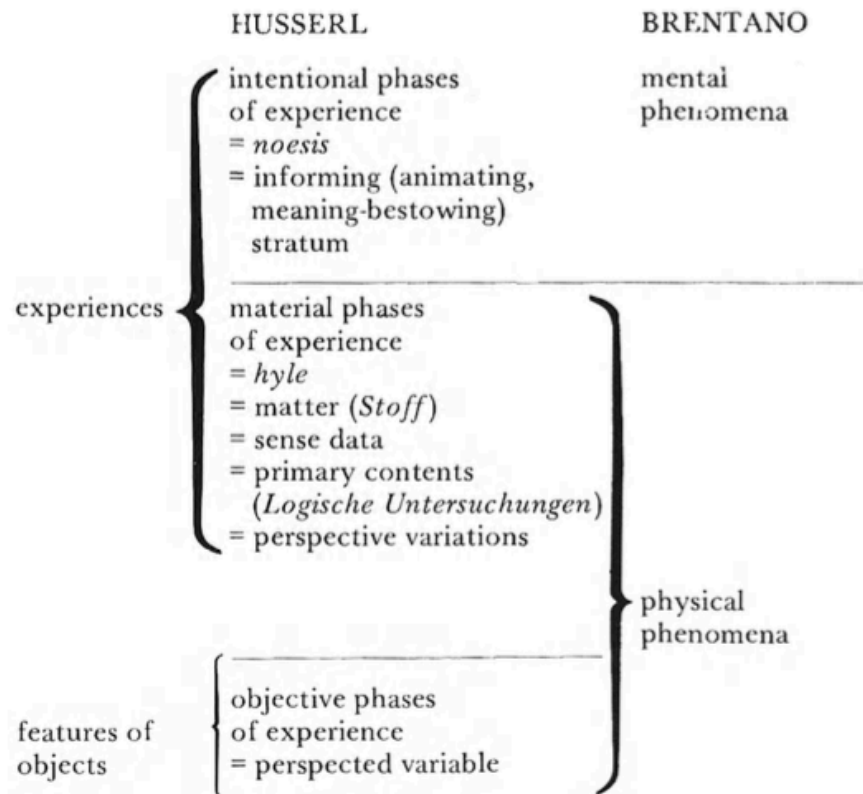
fregeano de Føllesdal y Dreyfus, el sentido y la referencia. Como bien leemos, para aclarar este asunto en Fernández Moreno (2009, 32), aunque ambas se articulan, no se pueden confundir: "La noción de sentido es así una noción de carácter *epistémico*; el sentido de una expresión es un rasgo de la expresión que conocemos cuando la entendemos". El sentido tiene que ver con este "objeto de pensamiento", del que nos habla Føllesdal, en aquellas determinaciones epistémicas en el tratamiento mental del objeto. Sin embargo, este objeto mental está en la determinación interna del físico real, por cuanto "la noción de sentido es asimismo una noción de carácter *semántico*, pues el sentido de una expresión *determina* su *referencia* o, como suele decirse, es el modo como la expresión nos presenta su referente" (Fernández Moreno 2009, *Ibíd.*). Este es el punto fregeano en Føllesdal, si bien este autor nos aclara, como podría bien corroborar Dreyfus, que Brentano y Husserl no son equiparables totalmente, en el austríaco la direccionalidad no depende tanto del contenido del objeto físico-real y de pensamiento sino a la estructura determinante de cualquier contenido. No es tanto una correlación sujeto-objeto, cuanto sujeto-noema-objeto:

By help of a trichotomy between the acting subject, the noema, and the object, Husserl is able to both retain Brentano's idea of the directedness of mental phenomena and overcome the difficulties where there are no corresponding objects. While Brentano kept insisting that the directedness of the act should be accounted for by means of some object toward which the act is directed, and got into all his problems, Husserl's view is that the directedness of the act should be accounted for not by some object toward which the act is directed, but by a certain structure of our consciousness when we are performing an act. This structure Husserl calls the noema. (Føllesdal 1982, 36)



La tricotomía de Husserl es preferible a la dicotomía de Brentano puesto que la dirrecionalidad mental no está subordinada en el austríaco al objeto, sea real, imaginado o ideal, sino al acto estructural del dirigirnos al objeto, una "cierta estructura" noemáticamente conforma el objeto. El objeto no es, en consecuencia, el que capitanea a lo que tiende la mente sino más bien es la mente estructurante lo que nos ha de concernir primeramente. Así, "that is the little point that makes all the difference between Brentano and Husserl", en la forma en la que la mente compone y estructura de manera unificada la tricotomía sujeto-noema-objeto, cuyos tres componentes nos pueden parecer heterogéneos solamente al precio de no ver el papel que tiene la mente al estructurar las cosas mismas en su sentido y referencia.

Empero, y dándole la razón al crítico Shim (2005), en Føllesdal (1982, 40), los *Empfindungsdaten* o datos sensoriales de Husserl se acaban por conceptualizar, cuando más bien en Husserl el componente hilético, si bien está ligado cercanamente al morfo-noético no puede reducirse a simple objeto mental. Føllesdal confunde (1982, *Ibíd.*) lo hilético "*according to Husserl, is that there is some object that impinges causally upon our sensory organs*", como si fuese un superestrato de sentido no claramente distinguido de lo conceptual. La mente en Husserl se hace así parecida a la estructura programada de los ordenadores que procesan información de múltiples formas: Husserl sigue siendo muy metalista a este respecto dado que es la mente la que decide todo. Pasamos, a continuación a ver a tenor de esto, en el siguiente esquema (Føllesdal 1982, 38), algunas distinciones de la teoría percepción tanto en Husserl como en Brentano.



**Figura 2**

En este esquema tan mentalista se puede apreciar todo lo que se puede ganar con Husserl a diferencia de Brentano, lo cual puede entenderse en sus respectivas teorías de la percepción. Mientras que en Brentano, la intencionalidad comprende los fenómenos mentales y los fenómenos físicos, cobrando importancia según su el objeto al que se dirige la intencionalidad mental, en Husserl la experiencias y las características de objetos forman parte de la misma estructura de la mente. Lo valioso es que esta estructura consta de fases intencionales de experiencia (lo que sería la *noesis*), en tanto lo más informacional y significacional de esta experiencia, mientras que las fases materiales de la experiencia –*hylé*– contiene su materia, datos sensoriales, contenidos primarios y variaciones de perspectivas. Lo más polémico en Føllesdal, que también puede imputarse a Dreyfus, es que esta fase material no acaba de distinguirse de lo que no pueda conceptualizarse. Se ve bien en el anterior esquema que, en Husserl, la materia, aunque es

experiencia moldea dicha experiencia. Según cómo se puedan variar las perspectivas de las cualidades de los objetos, podemos contar con las fases objetivas de la experiencia. Además, ni Føllesdal ni Dreyfus se dan cuenta de que, en Husserl, experiencia (*Erfahrung*) y conciencia (*Bewusstsein*) no se resumen en una suerte de internismo de conciencia, sino que la experiencia cubre los aspectos somáticos y sensorio-motrices; mientras que la conciencia puede ser empírica y también pura según su formalidad –esto cual nos conduce, como indicamos más arriba, a la Inteligencia Artificial de tipo formalista–.

Habida cuenta de esto, pensamos que Brentano no estaba tan desorientado, a tenor de las modernas fenomenologías de la corporalidad, al incluir en los fenómenos físicos las fases materiales de la experiencia y la de los objetos, aunque erró al aducir que “*physical phenomena exist only intentionally*” (Føllesdal 1982, *Ibíd.*), mientras que en el Husserl dreyfusiano la intencionalidad viene circunvalada por las fases hiléticas y objetivas de la experiencia. Hay, por tanto, una clara dubitación entre un Husserl trasmutado en Frege y otro que incluye lo somático y pre-conceptual en su fenomenología. No han faltado quienes, como Mohanty (1982, 52-53), ponen en duda a este Husserl fregeano de manera anacrónica en Føllesdal y Dreyfus ya que “(1) *Husserl arrived at the Vorstellung-meaning-reference distinction independently of Frege. (2) Husserl’s overcoming of psychologism and acceptance of a theory of objective pure logic was fundamentally independent of Frege’s 1894 review of Philosophic der Arithmetik. The basic change had occurred in 1891*”. En efecto, Husserl llegó a la distinción que, en vez del lenguaje fregeano, es entre *hylé* y *morphé* de manera independiente de Frege y, además de esto, su anti-psicologismo fue anterior –con un margen de tres años– a la obra señera del matemático. Esto es algo que Dreyfus ha debido avistar mejor, a la hora de seguir a pie juntillas a su maestro en fenomenología husserliana. Hoy son muchos los que someten a inclemente crítica la visión dreyfusiana de Husserl, como

puede ser el neurobiólogo y fenomenólogo chileno Francisco J. Varela (1996)<sup>3</sup>, ya que parecería que la mente computa de alguna forma, bajo reglas estructurales y estructurantes los *inputs* del mundo.

#### **4. La traslación de la fenomenología husserliana a la heideggeriana a favor de la Inteligencia Artificial**

Tras haber comprobado Dreyfus que la fenomenología de Husserl, siempre contemplándola con el antifaz de Føllesdal, sigue siendo demasiado mentalista y convierte a la mente en un conjunto de estructuras y reglajes de orden superior, el profesor de Berkeley intentará conquistar distinción y comparación en la postura del austríaco para con la fenomenología del alemán Martin Heidegger<sup>4</sup>. De hecho, John Searle (2000) ha dicho calificado a Dreyfus como un de los mejores comentadores de este filósofo<sup>5</sup>. Desde principios de la década de los 60, con su tesis de doctorado, Dreyfus (1963) propuso ir avanzando, según el subtítulo de la misma "*from transcendental to existential phenomenology*". A Husserl le costó dar un espacio y una importancia contundentes a la corporalidad, la situacionalidad y las, *verbi gratia*, –por seguir siendo más claros– necesidades cenestésicas–

Ciertamente, Heidegger no cubre estos tres desempeños a la perfección – el papel del cuerpo en él es demasiado espacial–, pero permite a la

---

<sup>3</sup> "Los mejores científicos cognitivos deberían haber leído la colección editada por Dreyfus, que presenta a Husserl como alguna tipo de proto-computacionalista, asumiendo que este pedacito de la historia es lo único que hay que saber sobre fenomenología. Esto se han convertido en una interpretación frecuentemente citada, si bien los críticos han destacado que la lectura cognitiva que hace Dreyfus de Husserl es seriamente fallida" (Varela 1996, 346).

<sup>4</sup> Nuestro autor ha publicado compendios sobre fenomenología y existencialismo (Cfr. Dreyfus y Wrathall 2006), así como otro expresamente sobre Heidegger (Cfr. Dreyfus y Wrathall 2006).

<sup>5</sup> Así lo reconocía el celeberrimo pensador: "Dreyfus has probably done more than any other English- speaking commentator to make the work of Heidegger intelligible to English- speaking philosophers" (2000, 71).

Inteligencia Artificial perfilarse en el aspecto de la existencia y no tanto, como en Husserl, en el de los superestratos y superestructuras mentales, los cuales desafortunadamente hacen aumentar la inversión investigadora en máquinas superinteligentes –entiéndasenos bien– que destaquen en inteligencia formal y no sentiente. Dreyfus consintió algunas de las reservas de Heidegger hacia la tecnología, si bien –aclarémoslo– nuestro autor no tuvo una postura tecnófoba sino que actuó contra una Inteligencia Artificial cargada de presupuestos y dogmatismo, que obstruían la consecución del objetivo de emular una inteligencia, si no igual, superior a la humana:

Thus, for Heidegger, technology, with its insistence on the "thorough going calculability of objects", is the inevitable culmination of the exclusive metaphysics, concern with beings (objects) and the concomitant exclusion of Being (very roughly our sense of the human situation which determines what is to count as an object). (Dreyfus 1972, 124)

El problema es que, en el alemán, la tecnología corona el pensamiento metafísico –en sentido denostado–, al retraerse en las cosas y no en el Ser. Este es el problema del mito de creer que una rama del saber manipuladora de objetos –de seres– puede llegar a atrapar el quicio de la totalidad –el Ser–. Sin embargo, Heidegger desplaza la fenomenología mentalista de Husserl a aquello que permite desde el exterior toda inteligencia. No es la inteligencia ni la teoría del conocimiento o de la percepción su mayor destreza y línea de investigación. Dreyfus (1972, 139) es de la opinión que Heidegger puede ser de gran interés para ver "*the limitations of technology*", pero no la desaparición de la misma<sup>6</sup>. Para dar todavía más

---

<sup>6</sup> Esto no es óbice para que Dreyfus reconozca el valor de la ontología del arte en Heidegger, aunque siempre la contraponga a su minusvaloración de la tecnología: "Heidegger's basic insight is that the work of art not only manifests the style of the culture; it articulates it. For everyday practices to give us a shared world, and so give meaning to our lives, they must be focused and held up to the practitioners. Works of art, when performing this function, are not merely representations of a pre-existing state of affairs, but actually produce a shared understanding" (Dreyfus 2005, 409). La diferencia de la tecnología y el arte, como acabamos de leer, es que la tecnología, movida por una razón

resonancia a su voz en lo que atañe al tema de este TFM, Heidegger testimonia las limitaciones de una tecnología de la Inteligencia Artificial, según los dos estadios diacrónicos –ya examinados por nosotros– emparentados con el mentalismo husserliano.

En Husserl, todo se subordina a las estructuras y superestratos de la mente que tiene conciencia experiencial, sin encajar bien en estas actividades mentales superiores un mundo que queda como subsuelo de menor interés. En Heidegger, por el contrario, el hombre no ve al mundo de manera aliena sino como su casa desde donde interpreta su existencia y los atisbos del Ser. Heidegger enlaza, en la misma analítica existencial del *Dasein*, mente-existencia-tiempo-historia-mundo, mientras que la más novata Inteligencia Artificial, por su parte, se encajeta en la mente. En *Ser y tiempo*, el hombre unifica los objetos del mundo a partir de su existencia en el mundo y no fuera de él en una mente sintetizadora o purificadora de las vaguedades gnoseológicas sobre dichos objetos:

In *Being and Time*, Heidegger gives a description of the human world in which man is at home, on the model of a constellation of implements (*Zuege*), each referring to each other, to the whole workshop and ultimately to human purposes and goals. (Dreyfus 1972, 173).

Todo lo que acaece en el mundo concierne a las intenciones y necesidades del ser humano del que es su casa: "*Heidegger tries to account for the pervasive concern organizing human experience in terms of a basic human need to understand one's being*" (Dreyfus 1972, 188). En Husserl, es la necesidad de entender y fundamentar la experiencia propia lo que contaba, mientras que en Heidegger es el lugar propio en el mundo, lo cual es una

---

instrumental y representacional, manipula objetos mentales o artefactuales, mientras que el arte es capaz de producir obras nuevas más del lado del Ser que de los entes: es una comprensión, en el seno de la cultura creativa, del Ser.

aportación –si lo llevamos al lado de la Inteligencia Artificial– a favor de unas máquinas inteligentes que pudiesen ser dotadas de una mayor adaptabilidad en el mundo. Pero Dreyfus no sigue a Heidegger contra Husserl a pie juntillas –lo cual es sugerente para nuestro tema sobre la relación entre la Inteligencia Artificial y la fenomenología Husserliana en su obra– sino que también critica las pocas destrezas prácticas de la analítica existencial del alemán.

De su filosofía de la existencia, nos señala Dreyfus que "*this analysis remains very abstract. It accounts for significance in general but not for any specific goal or specific significance*" (Dreyfus 1972, *Ibíd.*). Filosofar sobre la existencia y el mundo en general repercute negativamente en la poca utilidad de esto para la ingeniería en Inteligencia Artificial, sabiendo bien que la utilidad, vinculada a los utensilios, no era un asunto muy preocupante para Heidegger. Aún habiendo criticado la cibernética como ejemplo de la modernidad metafísica, a este pensador no le avisaba con tanta fuerza las soluciones y taxonomización útil a las preguntas cuanto la pregunta más radical por el sentido del Ser en la propia existencia mundanal: "*For Heidegger there can be no list of specifications which the solution must fulfill*" (Dreyfus 1972, *Ibíd.*). Se podría que Husserl era más especificador que su alumno pero, por desgracia, siempre protegido en el búnker mental que busca los sentidos de la referencia –si caricaturizamos de alguna manera a Føllesdal–.

Habida cuenta de esto, lo que propone Dreyfus, para acabar con una Inteligencia Artificial empecinada en superinteligencias formales, es una suerte de eclecticismo entre la fenomenología de Husserl –en las competencias mentales–, la de Heidegger –en las habilidades interactuantes en el mundo– y de fenomenólogos como Merleau-Ponty –en las destrezas del cuerpo y sus necesidades–. Por lo tanto, para comprender lo que nuestro autor nos brinda "*we require a more concrete*

*phenomenological analysis of human needs*". Esto podría inspirar a que los tecnólogos de la Inteligencia Artificial se planteen apartar de su precomprensión las aserciones biologicistas, psicologistas, epistemologistas y ontológicas que les impiden inventar máquinas con una inteligencia – permítasenos la redundancia– no-maquinista sino –y sin perjuicio de los altos niveles de cálculo y raciocinio a conseguir– que sea corporal, situacional y vitalmente intencional. Dreyfus pensaba que entender la intencionalidad sin lo demás, es decir, sin los aspectos corporalizados y situacionales fue un empobrecimiento del husserlianismo. Dreyfus (1995, ix) recalca que Heidegger, en oposición a Husserl, enfatiza "*the non-mentalistic approach to intentionality in Being and Time*". De hecho, nuestro autor osa afirmar que el primero pergeñó su fenomenología hermenéutica en esa obra para ir directamente contra su mentor:

Heidegger developed his *hermeneutic* phenomenology in opposition to Husserl's *transcendental* phenomenology. Husserl had reacted to an earlier crisis in the foundations of the human sciences by arguing that the human sciences failed because they did not take into account intentionality —the way the individual mind is directed at objects by virtue of some mental content that represents them. (Dreyfus 1995, 2)

El Husserl dreyfusiano apoya la visión de Føllesdal, por la que la única manera que se podía valorar al austríaco era desde una intencionalidad de la mente solitaria dirigida a objetos. Esto podría hacer a más de un exegeta vincular a Husserl con una Inteligencia Artificial centrada en la implementación y sobredimensión de las capacidades mentales y no de las afectivas e interactuantes en las máquinas. El ordenador que no trata contenido semántico vivido sino símbolos podría ser una consecuencia de esto, por cuanto que Husserl "*developed an account of man as essentially a consciousness with self-contained meanings, which he called intentional content*" (Dreyfus 1995, *Ibíd.*). El ser humano no sería en el austríaco sino



esencialmente conciencia auto-contenida, entregada enteramente a los contenidos intencionales. Dreyfus ha sido todavía más duro con Husserl, sosteniendo que lo suyo es un "*self-sufficient individual subject directed at the world by means of its mental content*" (Dreyfus 1995, 3), una acusación de autosuficiencia que revierte en la misma autosuficiencia de las destrezas de orden superior en las Inteligencia Artificial. A propósito de esto, Heidegger supone una ruptura para con Husserl que hemos de atender en este TFM si queremos saber, aunque sea *a contrario*, lo que vertebró la Inteligencia Artificial y la fenomenología husserliana.

Por ello, el ejemplo de Heidegger, pupilo inmediato del austríaco, nos ha sido magnífico. Dreyfus (Dreyfus 1995, *Ibíd.*) nos asegura que "*Heidegger breaks with Husserl and the Cartesian tradition*", porque lo suyo es más bien una fenomenología de la vida cotidiana, como tendría que ser la Inteligencia Artificial, es decir, una empresa de la vida corriente y, por ende, más compleja que el solo raciocinio. La firma de Heidegger ha sido desmentalizar la fenomenología, tal y como Dreyfus propone desmentalizar la Inteligencia Artificial: "*At the foundation of Heidegger's new approach is a phenomenology of 'mindless' everyday coping skills as the basis of all intelligibility*" (Dreyfus 1995, *Ibíd.*). Esta fenomenología –eventualmente trasladable a la Inteligencia Artificial– desmentalizada a favor de un adentrarse cotidiano en el mundo es lo deseado por nuestro autor. En el lado más alejado de Husserl, Heidegger "*seeks to minimize the role of the conscious subject in his analysis of human being*" (Dreyfus 1995, *Ibíd.*), si bien nuestro autor está convencido de la importancia de la dimensión consciente en la Inteligencia Artificial, aunque –como venimos insistiendo– implementándola con la corporalidad y la situacionalidad.

## 5. Conclusiones críticas

Para dar fin a nuestro TFM y tras estos cuatro puntos no podemos sino coincidir, parafraseando Searle (2000, 92), que "*Dreyfus has made many valuable contributions to philosophy*", especialmente, según la hipótesis principal que hemos enunciado en la Introducción, entre fenomenología-ciencia / fenomenología-tecnología, pero sobre todo en lo que concierne a la fenomenología husserliana, leída y revisada por Dreyfus, y la Inteligencia Artificial.

Por falta de tiempo y espacio, no hemos tenido ocasión de tratar aquí nada que tenga que ver con la manera en la que Dreyfus introduce el cognitivismo y el computacionalismo en la fenomenología *et viceversa*, lo cual sería de gran actualidad, aunque –reconozcamos la limitación de un TFM– lo dejamos para un trabajo posterior. Llegamos a la conclusión de que Dreyfus ha hecho apreciar que, para una más eficiente y eficaz Inteligencia Artificial, es necesario que cambie su proyecto primero de simulación cognitiva y procesamiento semántico, es decir, no tanto en la sombra husserliana mentalista sino heideggeriana, de un ser-en-el-mundo, o, como se ha tildado en la filosofía norteamericana, del "sentido común". En una obra dedicada a una filosofía del internet, el Dreyfus más reciente corrobora nuestras conclusiones críticas:

But in the late 1970s and early 1980s AI researchers reluctantly came to recognize that, in order to produce artificial intelligence, they would have to make explicit and organize the commonsense knowledge people share, and that was a huge task. (Dreyfus 2009, 17)

Sin este conocimiento del sentido común, las máquinas y programas inteligentes podrán hacer cualquier tarea formalmente magnífica menos habérselas con el mundo en las diversas prácticas sociales y comunicativas

cotidianas. Podríamos tachar, con Dreyfus (2009, 17) de "*failure of Artificial Intelligence*" el hecho de continuar, lo sepan o no sus inventores, con una fenomenología de tipo mentalista –que es probable que nuestro autor haya exagerado en Husserl motivado por Føllesdal–. Nuestro autor piensa que, siendo todavía obtusamente abstracta y maliciosamente cuestionadora de la tecnología, la fenomenología heideggeriana puede ser mucho más estimulante e iluminadora para una Inteligencia Artificial de máquinas que puedan al menos aproximarse en algo –dudamos seriamente si esto pueda darse de verdad algún día– a los seres humanos racionales y también sentientes, que puedan hacer metáforas, captar y emitir ironías, así como tener sentido del humor, aprender de los errores o crear cultura, a veces de manera revolucionaria y en ruptura con respecto a fases de la historia personal de cada uno o de las sociedades. Que los inventos de la Inteligencia Artificial puedan, en definitiva, sondear y adentrarse (*cope, coping*) en el mundo, supondría romper las barreras artificial-natural, lo cual es, a todas luces –y negarlo sería un absurdo–, con lo que contamos.

Toda esta complejidad humana es tremendamente difícil de lograr en programas y máquinas artificiales. Nuestra diferencia, por ahora insalvable, es que somos seres corporalizados, lo cual significa que somos constitutivamente abiertos y receptivos, en constante transformación y desarrollo, sumergidos en la vida: somos sistemas biológicos constantemente dinámicos y podemos reformar –o deformar– la realidad, que experimentamos al vivirla corporalmente. He aquí el fracaso de una Inteligencia Artificial imbuida de mentalismo y formalismo: "*The failure of AI projects [...] should call our attention to how important our bodies are in making sense of the world*" (Dreyfus 2009, 19). Por ello, finalmente, han sido tan importantes los tres desempeños que Dreyfus lleva remachando desde los años 70: el desempeño corporal, el situacional y el intencional –no sólo en lo privado sino lo colectivo, lo social y lo histórico–, todos ellos más propios de una fenomenología de la carne, que solamente de una

fenomenología de la conciencia, como la de Husserl. Creemos, en suma, que en este trabajo se ha visto la necesidad de las dos: tanto las capacidades mentales de orden superior como las del cuerpo en el mundo son muy incitantes para otro tipo –si algún día se obtiene cabalmente– de Inteligencia Artificial.

### **Bibliografía**

DREYFUS, H.L. (1972). *What computers can't do. A critique of Artificial Intelligence*, New York: Harper & Row.

\_\_\_ (1995). *Being-in-the-world. A commentary on Heideggers's Being and Time. Division I*, Cambridge: MIT Press.

\_\_\_ (1963). *Husserl's phenomenology of perception: from transcendental to existential phenomenology*, Tesis doctoral, Harvard University.

\_\_\_ (2005). "The work of art as articulating a culture's understanding of Being", en: H.L. Dreyfus y M. Wrathall, *A companion to Heidegger*, Oxford: Blackwell.

\_\_\_ (2009). *On the internet*, London: Routledge.

DREYFUS, H.L. y HALL, H. (1982), *Husserl, intentionality, and Cognitive Science*, Cambridge: MIT Press.

DREYFUS, H. y RABINOW, P. (1983). *Michel Foucault: Beyond structuralism and hermeneutics*, Cambridge: MIT Press.

DREYFUS, H.L. y WRATHALL, M. (2005). *A companion to Heidegger*, Oxford: Blackwell.

\_\_\_\_ (2006). *A companion to phenomenology and existentialism*, Oxford: Blackwell.

GÓMEZ ROMERO, I. (1995). "La teoría husserliana de la constitución en *Ideas I'*", *Anales del Seminario de Historia de la Filosofía*, 12, pp. 91-106.

FERNÁNDEZ MORENO, L. (2009). "Sentido, referencia y representación lingüística en Frege", *Contrastes*, 14, pp. 31-48.

FØLLESDAL, D. (1982). "Brentano and Husserl on Intentional Objects and Perception", en: H. Dreyfus y H. Hall, (eds.), *Husserl, intentionality, and cognitive Science*, Cambridge: MIT Press, pp. 31-41.

HEIDEGGER, M. (2005). *Ser y tiempo*, Santiago de Chile: Editorial Universitaria.

HUSSERL, E. (2008). *La crisis de las ciencias europeas y la fenomenología trascendental*, Buenos Aires: Prometeo Libros.

KUHN, Th. (1962). *The structure of scientific revolutions*, Chicago: University of Chicago Press.

MINSKY, M. (1986). *The society of mind*, New York: Simon & Schuster.

MOHANTY, J.N. (1982). "Husserl and Frege: A new look at their relationship", en: H. Dreyfus y H. Hall, (eds.), *Husserl, intentionality, and cognitive Science*, Cambridge: MIT Press, pp. 43-58.

RABANAQUE, L.R. (2010). "Actitud natural y actitud fenomenológica", *Sapientia*, 67, pp. 229-230.

SEARLE, J. (2000). "The limits of phenomenology", en: M. Wrathall y J. Malpas, *Heidegger, coping, and Cognitive Science. Essays in Honor of Hubert L. Dreyfus, Volume 2*, Cambridge: MIT Press.

SIMON, H. (1965). *The shape of automation for men and management*, New York: Harper & Row.

SHAFFER, C.A. (1998). *A practical introduction to data structures and algorithm analysis*, New Jersey, Prentice Hall.

SHIM, M. (2005). "The duality of non-conceptual content in Husserl's phenomenology of perception", *Phenomenology and the cognitive sciences*, 4, 2, pp. 209-229.

SMITH, A.D. (2003). *Husserl and the Cartesian Meditations*, New York: Routledge

VARELA, F.J. (1996). "Neurophenomenology: A methodological remedy for the hard problem", *Journal of Consciousness Studies*, 3, pp. 330-349.

WRATHALL, M. y MALPAS, J. (2000a). *Heidegger, authenticity and modernity. Essays in Honor of Hubert L. Dreyfus, Volume 1*, Cambridge: MIT Press.

\_\_\_\_ (2000b). *Heidegger, Coping, and Cognitive Science. Essays in Honor of Hubert L. Dreyfus, Volume 2*, Cambridge: MIT Press.

ZUBIRI, X. (1980). *Inteligencia sentiente. Inteligencia y realidad*, Madrid: Alianza Editorial.